STATISTICAL ANALYSIS PLAN

Efficacy of Hydroxychloroquine and Hydroxychloroquine + Azithromycin for Early Treatment of SARS-CoV-2 Infection for Prevention of COVID-19 Disease

Protocol Registry Number: NCT04354428

Effective Date: 24/01/2022

Version: 1.1

STATISTICAL ANALYSIS PLAN

Protocol Name:	Efficacy of Novel Agents for Treatment of SARS-CoV-2 Infection Among High-Risk Outpatient Adults: An Adaptive Randomized Platform Trial
Author(s):	Elizabeth R. Brown, ScD
Version:	1.1

Author(s):

Elizabeth R. Brown, ScD		24/01/2022
Protocol Statistician	Signature	Date: DD/MM/YYYY

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS

A list of abbreviations used in the SAP.

Term/Abbreviation	Definition
HCQ	Hydroxychloroquine
HCQ+A	hydroxychloroquine + azithromycin
LRTI	Lower respiratory tract infection
WHO	World Health Organization

1. INTRODUCTION

This version of the SAP focuses on the three arms first implemented in the protocol.

1.1. General Design Considerations

1.2. Study Objectives and Endpoints

The protocol was written early in the pandemic when there was not adequate information about typical disease course. Between then and the writing of this document, our understanding of COVID-19 disease progression has changed, resulting in changes to some of the objectives and endpoints. Most notably, the trial was stopped early for operational futility on the primary clinical objectives. These changes are noted in the table below.

Table 1 Original objectives and endpoints as written in the protocol

Objectives	Endpoints
Primary	
To test the efficacy of HCQ and HCQ + azithromycin compared to placebo to prevent progression to LRTI, among persons with SARS-CoV-2 infection who are at high risk of progression	LRTI, defined by resting SpO2<93%, sustained for 2 readings 2 hours apart AND presence of subjective dyspnea or cough
To test the efficacy of HCQ and HCQ + azithromycin compared to placebo to prevent hospitalization and death measured at Day 28 among participants who are at high risk of progression	Cumulative incidence of hospitalization or mortality, measured at Day 28
To test the efficacy of HCQ and HCQ + azithromycin compared to placebo to reduce SARS-CoV-2 viral shedding	Time to clearance of nasal SARS-CoV-2, defined as 2 consecutive negative swabs (swabs collected daily Day 1 through Day 14)
Secondary	
1. To test the safety of HCQ and HCQ + azithromycin compared to placebo for	Serious adverse events (including death and hospitalization) and adverse events resulting in treatment discontinuation

treatment of high-risk outpatients with SARS-CoV-2 infection	
2. To test whether HCQ and HCQ + azithromycin has an effect on the duration of hospitalization among persons who become hospitalized with COVID-19 disease	Days of hospitalization
3. To test whether HCQ and HCQ + azithromycin has an effect on disease severity compared to placebo	• Peak score on WHO Ordinal Scale for Clinical Improvement • Peak score on modified Flu-PRO within the first 14 days
4. To test whether HCQ and HCQ + azithromycin decreases the resolution rate for symptomatic SARS-CoV-2 infection/COVID-19 disease compared to placebo	• Proportion of days with fever after randomization • Proportion of days with respiratory symptoms after randomization • Proportion of days with SpO2<93% without supplemental oxygen after randomization
5. To test whether HCQ and HCQ + azithromycin is associated with decreased viral shedding from self-collected nasal swabs over 14 days compared to placebo	• Proportion of days with SARS-CoV-2 detected from mid-nasal swabs by PCR • Quantity of SARS-CoV-2 detected from mid-nasal swabs by PCR

1.3. UPDATED OBJECTIVES AND ENDPOINTS

The trial was stopped early due to operational futility the meet the first two primary objectives (development of LRTI and hospitalization/death). In light of this several of the secondary endpoints have been updated as follows.

- Because of the very few number of hospitalizations, the Secondary Objective 2 will no longer be evaluated.
- The endpoint for Secondary Objective 3 will be the Flu-PRO score evaluated daily over 14 days instead of only the maximum value. Appropriate statistical methods will be detailed in Section 7.2.
- The CDC definition of COVID-19 will now define the endpoint for Secondary Objective (SO) 4. The endpoints currently presented in Table 1 for SO 4 will be summarized over time as described in Section 7.2.

1.4. RANDOMIZATION

Due to the adaptive design of the study, we implemented randomization to 12 dummy codes, stratified by site and risk cohort. Participants are randomized to 1 of 12 dummy codes that are mapped to active and control arms using an allocation ratio appropriate to the current design of the study. Participants were randomized 1:1:1 for the Hydroxychloroquine and Hydroxychloroquine + Azithromycin efficacy study.

1.5. SAMPLE SIZE AND POWER

When the protocol was written, we assumed that the trial would enroll quickly and therefore end a short time after starting, making interim analyses unattainable. Unfortunately, that has not been the case due to changes in the epidemic and controversy surrounding hydroxychloroquine. This SAP makes minor changes to the study design to allow for multiple interim looks at efficacy and futility and therefore includes minor changes to enrollment targets. The new sample size calculations are described here. All sample size calculations are based on 4 interim analyses with one final analysis and stopping rules for efficacy and futility.

1.5.1. LRTI and Hospitalization Endpoints

Overall sample size will depend on the number of events observed. To detect at least 50% treatment efficacy in reducing LRTI or hospitalization/death under a null hypothesis of no reduction, four interim monitoring analyses with stopping for futility and efficacy based on the O'Brien-Fleming bounds and 90% power, the trial is designed to achieve 93 events per pairwise comparison for the LRTI and hospitalization endpoints. For two active arms and one control arm, the total number of events required for each endpoint is 124.

As the initial sample size target, 165 per arm was chosen for each experimental group (HCQ, HCQ + azithromycin), assuming a control arm rate of approximately 30% and 5% dropout rate.

1.5.2. Viral Shedding Endpoint

Assuming the median time to cessation of viral shedding is 10 days, with a constant baseline hazard for cessation of shedding, 90% statistical power, 1-sided type I error rate of 2.5% for efficacy and early stopping for futility, the design (including interim analyses) can detect 60% or greater increase in the rate of shedding cessation in the combined high- and low-risk cohorts with 200 viral clearance events per pairwise comparison.

2. GENERAL DATA ANALYSIS CONSIDERATIONS

Descriptive statistics that will be used to summarize continuous variables are as follows: mean and standard deviation, median and interquartile range, quartiles, range, and number of missing data values. If applicable, continuous data that are skewed will be transformed to meet assumptions of normality. For categorical variables, descriptive statistics that will be used include the following: frequencies, relative frequencies, and the number of missing data values. Descriptive analyses summarizing baseline and follow-up data will be stratified by site and/or country. Line, scatter and box plots will be used, as appropriate, for longitudinal data representations. An alpha level of 0.05 will be used for all statistical tests unless otherwise specified.

2.1. BASELINE DATA

The following baseline variables will be described: age, sex, presence of known COVID-19 risk factors (includes all factors that categorize participants into the high-risk category), race/ethnicity and time since onset of symptoms.

Baseline variables will be summarized overall and within high and low risk strata by treatment arm. Because of the potential for imbalance between arms due to household-based randomization, statistical testing across arms will be performed using gee adjusted by risk strata with exchangeable correlation within household. If there appears to be imbalance across the arms, ($p<0.10$), the variable will be adjusted for in primary analyses.

2.2. Analysis Cohorts and Data Sets

This section describes the cohorts and corresponding datasets for the primary, secondary and exploratory analyses described in this document. In the definitions below, cohort refers to the study participants included in a dataset; whereas the datasets further specify which visits are included from the specified cohort. All data sets include the baseline variables specified in Section 2.1 and are included as guidance, recognizing that in practice, what is listed as a single data set may actually require sets of data sets (for example repeated vs single time point outcomes for secondary outcomes) depending on the analysis.

Intention to Treat (ITT) Cohort: The ITT Cohort includes all participants enrolled into the study.

Modified Intention to Treat (mITT) Cohort: The mITT Cohort includes all participants enrolled into the study who completed the day 1 (enrollment) questionnaire.

> ***Viral Shedding (VS) Data Set*** This data set will include all participants in the cohort with all data required for primary and secondary viral shedding endpoints.
>
> ***Hydroxychloroquine (HCQ) Secondary Efficacy Data Set*** This data set will include all participants in the cohort randomized to hydroxychloroquine or contemporaneous control and secondary endpoints.
>
> ***Hydroxychloroquine + azithromycin (HCQ+A) Secondary Efficacy Data Set*** This data set will include all participants in the cohort randomized to the combination hydroxychloroquine/azithromycin arm or its contemporaneous control and secondary endpoints.
>
> ***Hydroxychloroquine (HCQ) Safety Data Set*** This data set will include all participants in the cohort randomized to hydroxychloroquine or contemporaneous control and secondary outcomes.
>
> ***Hydroxychloroquine + azithromycin (HCQ+A) Safety Data Set*** This data set will include all participants in the cohort randomized to the combination hydroxychloroquine/azithromycin arm or its contemporaneous control and safety outcomes.
>
> **Hospitalized Data Set** This data set will include all participants who reported a COVID-19-related hospitalization within the planned 28 days of follow-up.

Modified Intention to Treat (ITT) High Risk (HR) Cohort: The ITT HR Cohort includes all mITT cohort participants enrolled into the study's high risk cohort.

> ***Hydroxychloroquine (HCQ) Efficacy Data Set*** This data set will include all participants in the cohort randomized to hydroxychloroquine or contemporaneous control.
>
> ***Hydroxychloroquine + azithromycin (HCQ+A) Efficacy Data Set*** This data set will include all participants in the cohort randomized to the combination hydroxychloroquine/azithromycin arm or its contemporaneous control.

Intention to Treat (ITT) Low Risk (LR) Cohort: The ITT LR Cohort includes all contacts enrolled into the study's low risk cohort (persons age 18-59 without risk factors for severe disease, including pulmonary disease, DM1 or DM2, HTN, immunocompromising condition, or BMI>30)).

Modified Intention to Treat (mITT) LRTI Cohort: The mITT LRTI cohort includes ITT HR Cohort participants who did not meet the criteria for the primary LRTI endpoint on day 1. The following data sets are created from participants in this cohort.

> ***Hydroxychloroquine (HCQ) LRTI Efficacy Data Set*** This data set will include all participants in the cohort randomized to hydroxychloroquine or contemporaneous control.
>
> ***Hydroxychloroquine + azithromycin (HCQ+A) LRTI Efficacy Data Set*** This data set will include all participants in the cohort randomized to the combination hydroxychloroquine/azithromycin arm or its contemporaneous control.

Modified Intention to Treat (mITT) Disease Resolution Cohort: This cohort includes mITT Cohort participants who meet the COVID-19 disease criteria on day 1 (Section 7.3). The following data sets are created from participants in this cohort.

> ***Hydroxychloroquine (HCQ) Disease Resolution Data Set*** This data set will include all participants in the cohort randomized to hydroxychloroquine or contemporaneous control.
>
> ***Hydroxychloroquine + azithromycin (HCQ+A) Disease Resolution Data Set*** This data set will include all participants in the cohort randomized to the combination hydroxychloroquine/azithromycin arm or its contemporaneous control.

Modified Intention to Treat (mITT) Viral Shedding Cohort: This cohort includes mITT Cohort participants who do not meet the definition for clearance of viral shedding at baseline (day 1 and day 2 negative swabs). The following data sets are created from participants in this cohort.

> ***Hydroxychloroquine (HCQ) Viral Shedding (VS) Data Set*** This data set will include all participants in the cohort randomized to hydroxychloroquine or contemporaneous control.
>
> ***Hydroxychloroquine + azithromycin (HCQ+A) Viral Shedding (VS) Data Set*** This data set will include all participants in the cohort randomized to the combination hydroxychloroquine/azithromycin arm or its contemporaneous control.

3. TRIAL PARTICIPANT DISPOSITION

3.1. DISPOSITION OF PARTICIPANTS

All of the participants who screened for and entered the study will be accounted for in a CONSORT flow diagram stratified by cohort (high vs. low risk) including the following measures:

The numbers of participants screened, screened out (with reasons for exclusion), enrolled, randomized, allocated to each study arm (with reasons for not receiving assigned intervention), reasons for discontinuation from follow-up (e.g., lost to follow-up, adverse event, noncompliance, etc.), and the number of participants with available data for each endpoint.

3.2. TREATMENT EXPOSURE

Self-reported pill taking will be reported by prescribed dosing day stratified by arm and overall. The summary measures will be number of doses taken and number of doses taken divided by

number prescribed. The number and proportion of compliant participants will be reported by day and overall time.

4. EFFICACY/EFFECTIVENESS ANALYSES

4.1. EFFICACY ANALYSES

All efficacy analyses will be adjusted for age, sex at birth and time from symptom onset to enrollment. These adjustments will be additive unless otherwise specified.

All analyses will account for the household clustering using robust standard errors as available. If not available, bootstrap methods will be used.

For time to event analyses, baseline (time 0) will be day of the enrollment visit.

Endpoints will be summarized as number of participants: hospitalized without prior LRTI, hospitalized with prior LRTI, died without prior LRTI, died with prior LRTI, who meet LRTI definition. These will be shown as counts and percentages overall and by cohorts defined in Section 2.2 within further breakdown by arm.

4.1.1. LRTI objective

Objective: To test the efficacy of HCQ and HCQ + azithromycin compared to placebo to prevent progression to LRTI, among persons with SARS-CoV-2 infection who are at high risk of progression.

Outcome: LRTI, defined by resting SpO2<93%, sustained for 2 readings 2 hours apart AND presence of subjective dyspnea or cough within the first 14 days of follow-up.

Cohorts: mITT LRTI Cohort

Data Sets: HCQ LRTI Efficacy Data Set, HCQ+A LRTI Efficacy Data Set

Definition of Endpoint: A participant will be defied as having the endpoint if they have two resting SpO2 readings less than 93% two hours apart but not more than 48 hours apart AND report subjective dyspnea (defined as a response of “somewhat or higher” on the “trouble breathing” question on flupro) or cough (defined as “somewhat or higher” as self-reported wet or dry cough) within 24 hours of the time the SpO2 is reported during in the first fourteen days of follow-up. The event time will be recorded as the day of the first SpO2 reading less than 93% with the presence of subjective dyspnea or cough. Participants who do not reach this endpoint will be censored at the last time that the endpoint was ascertained not to be present. Hospitalization or death due to COVID-19 may be a competing risk. Per FDA guidance, participants who are hospitalized or die due to COVID-19 will have an event time that is the first of LRTI, hospitalization or death.

Analysis Details:

Kaplan-Meier cumulative incidence curves will be estimated and plotted for each intervention including 95% confidence intervals accounting for household cluster.

Cox proportional hazards models stratified by site and adjusted for baseline covariates identified above will be used to estimate the efficacy as 1 minus the hazard ratio. Due to potential clustering induced by the household design, robust standard errors will be estimated for the hazard ratio. The overall efficacy will be tested using a p-value ascertained from the Wald statistic based on

the robust standard error. 95% confidence intervals will also be calculated based on the robust standard error. If there are too few endpoints to fit a Cox model then the proportion of participants meeting the endpoint in the first 14 days will be reported. Each active arm will be compared to the control arm separately using Fisher's Exact test.

4.1.2. Hospitalization/Death objective

Objective: To test the efficacy of HCQ, HCQ + azithromycin, or LPV/r compared to placebo to prevent hospitalization and death measured at Day 28 among participants who are at high risk of progression

Outcome: COVID-19 related hospitalization or death at or before day 28 or follow-up

Cohort: mITT High Risk Cohort

Data sets: HCQ Efficacy Data Set, HCQ+A Efficacy Data Set

Definition of Endpoint: A participant will be defined as having the endpoint if they were hospitalized or died due to COVID-19 as determined by the adjudication committee before day 28.

Analysis Details:

The proportion of participants with a hospitalization or death before the end of planned follow-up will be reported. Each active arm will be compared to the control arm separately using Fisher's Exact test.

4.1.3. Duration of Shedding Objective

Objective: To separately estimate and test the efficacy of HCQ and HCQ + azithromycin placebo to reduce SARS-CoV-2 viral shedding in high and low risk participants combined.

Outcome: Time to clearance of nasal SARS-CoV-2, defined as 2 consecutive negative swabs (swabs collected daily Day 1 through Day 14).

Cohort: mITT Viral Shedding Cohort

Data sets: HCQ VS Data Set, HCQ+A VS Data Set

Endpoint: The endpoint time will be set as the day of the first negative swab. Hospitalization and death are also competing risks for this outcome; however, clearance is a positive outcome whereas hospitalization and death are negative outcomes that may impede our ability to detect clearance. Participants who are hospitalized or die will have their event time censored at day 14. Other participants who do not have 2 consecutive negative tests will be censored on the day of their last positive test.

Analysis details:

Same as Section 4.1.1.

5. INTERIM MONITORING

A Data Monitoring Committee (DMC) will be convened to monitor the progress of the study. This will include study outcomes as well as futility and efficacy on the LRTI endpoint.

On July 23, 2020, the DSMB met and enrollment into the trial was stopped on July 28, 2020 for operational futility. No interim analyses were performed.

6. SAFETY ANALYSES

All safety analyses will be performed on the ITT cohort with outcomes presented by randomization arm reported overall and stratified by high and low risk.

6.1. QT PROLONGATION

To investigate whether the interventions lead to QT prolongation as measured by Corrected QT interval (QTc), we will create boxplots of QTc by study day and arm. To test for an overall difference between arms, we will fit linear mixed effects models of QTc post randomization through day 14 with random intercepts and time slopes. Randomization arm will be a main effect with additional main effect adjustments for time, age, sex, site and cohort.

6.2. ADVERSE EVENTS AND DEATHS

Serious adverse events (SAEs, including hospitalization and death), reportable adverse events leading to treatment discontinuation (RAEs), and cardiac v. non-cardiac events will be summarized in frequency tables (including both the number of each SAE/RAE and the number of distinct participants ever reporting each SAE/RAE) by severity and randomization arm, for each of the low and high risk cohorts, and overall.

7. NON-SAFETY SECONDARY ENDPOINTS

7.1. HOSPITALIZATION

Objective: To test whether HCQ or HCQ + azithromycin has an effect on the duration of hospitalization among persons who become hospitalized with COVID-19 disease compared to control.

Outcome: Days of hospitalization

Cohort: ITT Cohort

Data sets: Hospitalized Data Set

Endpoint: Number of days from admission to discharge (this may be informatively censored by death)

Analysis details:

Because of very low incidence of hospitalization in this study, only the minimum, median and maximum of days of hospitalization will be reported by arm.

7.2. DISEASE SEVERITY

Objective: To test whether HCQ or HCQ + azithromycin has an effect on disease severity compared to placebo

Outcomes: Disease severity is measured using two scales: the WHO Ordinal Scale for Clinical Improvement and The Modified Flu-PRO

Cohort: ITT Cohort

Data sets:

Endpoints: This analysis will have two endpoints: 1) the WHO Ordinal Scale for Clinical Improvement measured on days 14, 21 and 28 and 2) the score from the modified Flu-PRO measured daily on the first 14 days in the study, excluding the baseline (day 1). The modified Flu-PRO score is calculated as the mean score across the Flu-PRO questionnaire (sum numeric value over all symptoms and divided by number of symptoms[1]).

Analysis details: The distribution of scores will be described by arm at each visit including minimum, maximum, median and IQR. Boxplots of the Flu-PRO score will be presented by visit and arm and further stratified by risk cohort.

Testing: Because the maximum observed value may have undesirable statistical properties, we will perform hypothesis testing on all the observed data over the first 14 days. Mixed effects models with random slopes and intercepts will be used to model the mean Flu-PRO score over time with a main effect of randomization. The interaction between randomization and time will also be included to explore the potential effect of randomization on the change in Flu-PRO score over time. Testing between arms will be based on the coefficients from the main treatment effect and the interaction term using Wald tests. These tests will detect average differences between the arms in Flu-PRO scores. Recognizing that linear models may not well fit the Flu-PRO trajectories, an exploratory model using B-splines will also be fit.

7.3. SYMPTOM RESOLUTION

Objective: To test whether HCQ or HCQ + azithromycin decreases the resolution rate for symptomatic SARS-CoV-2 infection/COVID-19 disease compared to placebo

Outcome: Symptomatic SARS-CoV-2 infection/COVID-19 disease is defined based on the following criteria:

- At least TWO of the following symptoms: Fever (≥ 38ºC), chills, rigors, myalgia, headache, sore throat, new olfactory and taste disorder(s), OR
- At least ONE of the following symptoms: cough, shortness of breath or difficulty breathing, OR
- Severe respiratory illness with at least 1 of the following:
    - Clinical or radiological evidence of pneumonia, OR
    - Acute respiratory distress syndrome (ARDS)

A participant will be defined as having reached the endpoint if they have two consecutive visits that do not meet the definition above. The time of the endpoint will be the midpoint between the last symptomatic day before reaching the endpoint and the first non-symptomatic day defining the endpoint. Participants who do not have the endpoint will be censored at the last time they were

known to be symptomatic. Death and hospitalization will be counted in the endpoint by assigning their follow-up to be censored at day 14.

Cohort: mITT Disease Resolution Cohort

Data sets: HCQ Disease Resolution Data Set, HCQ+A Disease Resolution Data Set

Endpoint: Resolution of COVID-19 symptoms for 2 or more days.

Stratification variables: Indicator of more than 4 days since onset of symptoms (as measured at enrollment visit)

Analysis details: Kaplan-Meier cumulative incidence curves will be estimated and plotted for each intervention including 95% confidence intervals accounting for household cluster. Curves will also be plotted by arm and stratification variable.

Cox proportional hazards models stratified by high/low risk and adjusted for baseline covariates identified above will be used to estimate the efficacy as 1 minus the hazard ratio. Due to potential clustering induced by the household design, robust standard errors will be estimated for the hazard ratio. The overall efficacy will be tested using a p-value ascertained from the Wald statistic based on the robust standard error. 95% confidence intervals will also be calculated based on the robust standard error.

7.4. MAGNITUDE OF VIRAL SHEDDING

Objective: To test whether HCQ or HCQ + azithromycin is associated with decreased viral shedding from self-collected nasal swabs over 14 days compared to placebo

Outcomes: Quantity of SARS-CoV-2 detected from mid-nasal swabs by PCR as measured by cycle threshold (Ct)

Cohort: mITT Cohort

Data sets: VS Data Set

Analysis details: Generalized linear mixed effects models with random intercepts and slopes (time) will be used to model the quantity of SARS-CoV-2 over time. Because of the high number of samples where Ct was above the upper limit (eg >40) and not recorded, we will do analyses with Ct dichotomized as follows:

- Ct>40 on both targets vs Ct<=40 on either target
- Concordant results vs discordant or negative
- Ct<24 on either target as this may be an indicator of greater propensity to be infectious[2].

Fixed effects will include randomization arm, B-spline coefficients for time, an interaction between arm and time, cohort, age, and other baseline covariates indicated for the primary analysis. Time will be the time since enrollment visit to collection of the sample.

8. FOLLOW-UP SUMMARIES

Additional descriptive analyses will be performed on the following variables as recorded on participant surveys over time: Temperature, O2 saturation, heart rate and respiratory rate. These analyses will include boxplots by visit and arm, further broken out by risk cohort and table reporting mean, median and IQR by visit and arm stratified by risk cohort.

9. REFERENCES

1. Powers, J.H., 3rd, et al., *Reliability, Validity, and Responsiveness of InFLUenza Patient-Reported Outcome (FLU-PRO(c)) Scores in Influenza-Positive Patients.* Value Health, 2018. **21**(2): p. 210-218.
2. Bullard, J., et al., *Predicting infectious SARS-CoV-2 from diagnostic samples.* Clin Infect Dis, 2020.

10. CHANGE HISTORY

Major changes after Version 1.0 will be recorded here.

Version		Affected Section(s)	Activity Description
Number	Effective Date		
1.1	24/01/2022	Title Page	Corrected Protocol Registry Number